524 Rec'd PCT/PTO 0 1 OCT 1999

FORM PTO-1390 (REV 11-98) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: 097037014006

U.S APPLICATION NO. (If known, see 37 CFR 1.5): 09/381899

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/SE98/00602	April 1, 1999	April 1, 1997

TITLE OF INVENTION: METHOD AND ARRANGEMENT FOR AUTOMATIC DATA ACQUISITION OF FORMS

APPLICANT(S) FOR DO/EO/US: Jan Andersson

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [X] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. [ ] This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. [ ] A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. [X] A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. [ ] is transmitted herewith (required only if not transmitted by the International Bureau).
   b. [X] has been transmitted by the International Bureau.
   c. [ ] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. [ ] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. [ ] are transmitted herewith (required only if not transmitted by the International Bureau).
   b. [X] have been transmitted by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [ ] have not been made and will not be made.
8. [ ] A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. [ ] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. [ ] A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. [ ] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [ ] A FIRST preliminary amendment.
    [ ] A SECOND or SUBSEQUENT preliminary amendment.
14. [X] A substitute specification.
15. [ ] A change of power of attorney and/or address letter.
16. [X] Other items or information: Published International Application WO 98/47098 with search report, International Preliminary Examination Report with amended claims and Request.

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as "Express Mail Post Office to Addressee" Mailing Label Number EI962924597US in an envelope addressed to: Assistant Commissioner for Patents, BOX/PCT/DO/EO/US, Washington, DC ~~20231~~ On October 1, 1999

Merri C. Merrill

U.S. APPLICATION NO (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO	ATTORNEY'S DOCKET NUMBER
09/381899	PCT/SE98/00602	097037014006

17. [X] The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

		CALCULATIONS PTO USE ONLY
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO	$970.00	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO	$840.00	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO	$760.00	
International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4)	$670.00	
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)	$96.00	

ENTER APPROPRIATE BASIC FEE AMOUNT =				$ 970.00	
Surcharge of **$130.00** for furnishing the oath or declaration later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	
CLAIMS	NUMBER FILED		NUMBER EXTRA	RATE	
Total claims	34	- 20 =	14	X $18.00	$ 252.00
Independent claims	2	- 3 =	0	X $78.00	$ 0
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			0	+ $260.00	$ 0
TOTAL OF ABOVE CALCULATIONS =					$ 1,222.00
Reduction of 1/2 for filing by small entity, if applicable. A Small Entity Statement must also by filed (Note 37 CFR 1.9, 1.27, 1.28).					$ 611.00
SUBTOTAL =					$ 611.00
Processing fee of **$130.00** for furnishing the English translation later than [ ]20 [ ]30 months from the earliest claimed priority date (37 CFR 1.492(f)). +					$
TOTAL NATIONAL FEE =					$ 611.00
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +					$
TOTAL FEES ENCLOSED =					$ 611.00
				Amount to be: refunded	$
				charged	$

a. [X] A check in the amount of $506.00 to cover the above fees is enclosed.

b. [X] Please charge my Deposit Account No. 10-1202 in the amount of $105.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

c. [X] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 10-1202. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Jones, Day, Reavis & Pogue
77 West Wacker Drive
Chicago, IL 60601-1692

Phone: (312) 782-3939
Fax: (312) 782-8585

SIGNATURE:

Russell L. McIlwain
NAME

28,641
REGISTRATION NUMBER

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant or Patentee: ANDERSSON, Jan
Serial No.: 09/381,899
Filed or Issued: April 1, 1999
For: Method and Arrangement for Automatic Data Acquisition of Forms
(Related to: International Application No. PCT/SE98/00602; I.A. filing date 1 April 1998; priority date 1 April 1997)

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY STATUS
(37 CFR 1.9(f) and 1.27(c)) - SMALL BUSINESS CONCERN

I hereby declare that I am

[ ] the owner of the small business concern identified below:
[x] an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF CONCERN **ReadSoft AB**
ADDRESS OF CONCERN **Garnisionsgatan 25A**
SE-254 66 Helsingborg, Sweden

I hereby declare that the above identified small business concern qualifies as a small business concern as defined in 13 CFR 121.3-18, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees under §41(a) and (b) of Title 35, United States Code, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, the number of employees of the business concern is the average over the previous fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year; and concerns are affiliates of each other when either, directly or indirectly, one concern controls or has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the above entitled invention described in

[ ] the specification filed herewith
[X] application serial no. 09/381,899, filed April 1, 1999
[ ] patent no.:

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights to the invention is listed below and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b)).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING Jan Andersson
TITLE OF PERSON OTHER THAN OWNER Managing Director
ADDRESS OF PERSON SIGNING Garnisionsgatan 25A
SE-254 66 Helsingborg, Sweden

SIGNATURE _____ DATE 99.10.12

Title

Method and Arrangement for Automatic Data Acquisition of Forms

Technical field

The present invention refers to a method and arrangement for the automatic data acquisition, by means of a means for the same, of forms whose design and information is not known in advance, by input into the said means together with storage of patterns of the same.

The prior art

It is a problem for companies, organisations and others to make good use of the information found in different types of paper forms, documents, etc.

With new, modern technology, these items can be scanned with a scanner and entered into a database via commercially available software programs. However, sorting, identification and other checking routines must to a large extent still be performed manually via the computer's display or screen.

For example, to store an invoice from one and the same company as one specifically designed document with a logotype and other visual elements, it must be revised so that its format is adapted to one that can be accepted by the software and then scored in a database. This and other procedures must be repeated each time an invoice with a new design is scanned with the software.

To identity an invoice from a company that is already registered, the whole invoice is often searched, which is time consuming. Certain software programs can have search routines that restrict the extent of this searching. It is, however, difficult to safeguard against blurred or hand-written lines of text, etc.

A need therefore exists for all who handle invoices and other forms to quickly be able to identify these and/or quickly be able to enter and store new patterns in their invoicing system.

Patent US-A-4 933 979 describes traditional data acquisition from forms and requires pre-defined templates/patterns with no self-learning (adaptive) ability.

Patent US-A-5 140 650 mentions data acquisition from forms with what is known as "Form out" technology to cover-over the original document and only retain the parts that are "filled-in." This data acquisition is often combined with data acquisition according to US-A-4 933 979. The patent does not have any adaptive function for data acquisition of unknown forms.

Another patent, US-A-5 293 429, concerns the classification of documents with the help of lines on the documents and does not directly concern data acquisition or any adaptive function

for this. USA-A-5 293 429 does not ensure the identification of lines with object areas (areas with text) and a "RCG-value" (ReCoGnition, a number that uniquely identifies a document).

None of the said patents generates a form map for a form unknown to the system according to the patent and stores the map in real time in a form database for recognition at the next opportunity for identification. For the inventions according to these patents, the unknown form must therefore be stored later by other means.

Summary of the invention

One of the objectives of the present invention is to solve the problems named above as well as others during what is known as automatic data acquisition (interpretation) in connection with the handling of paper-based information.

The present invention concerns a system (method and arrangement) for the automatic data acquisition of forms where the system has no prior knowledge of what the form looks like or where on the form the information is to be found. In this way, templates of forms do not have to be defined in advance, but are instead registered as they are submitted to the system, i.e., in real time.

To accomplish the above objectives, the present invention specifies a method and arrangement for the automatic data acquisition, by means of a means for the same, of forms whose design and information is not known in advance, by input into the said means together with storage of patterns of the same. The method is adaptive, by which it includes learning and registering of forms as patterns without filled-in text, and by it also including the following steps for accomplishing the adaptive registration:

generation of a form map based on the previously unknown form's design for identifying information contained on the form;

searching and comparing the form map with stored, registered maps in a means for storing form maps;

storing generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

indication of agreement according to the limits for agreement when agreement is found; and

continued data acquisition for identifying of the information content of the form.

According to one embodiment of the present invention, the form map can consist of an object

area list with objects contained in the form whereby the object comprises colours and/or wholly or partly of text.

In an alternative embodiment, the form map constitutes a line map comprising objects in the form of coloured lines from the form.

Horizontal lines in the line map are used to produce a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

Vertical lines in the line map are used to produce a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

At least one line element that is included in a segment is marked in the equivalent key position, and segments that lack line elements remain unmarked in the equivalent key position.

A horizontal key and/or a vertical key constitute a line key in the line map, whereby during the said searching, the line key generated is compared with line keys stored in the means for verifying agreement.

The line keys are sorted in the storage means according to the number of markings.

The object's horizontal position in the object area list is used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

The object's vertical position in the object area list is used to produce a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

At least one object that is included in a segment is marked in the equivalent key position, and segments that lack objects remain unmarked in the equivalent key position.

A horizontal key and/or a vertical key constitute an object key in the object area list, whereby during the said searching, the object key generated is compared with object keys stored in the means for verifying agreement.

The object keys are preferably sorted in the storage means according to the number of markings.

Searching results in a pre-defined number of requested probable candidates for the currently searched form.

If needed, an operator can support manually the whole or parts of the adaptive registration or identification of the new form or registered forms respectively if several alternative candidates are found as probabilities according to a factor of merit.

Finally, the identity of the form is confirmed by the data acquisition of a RCG-value.

Furthermore, the present invention specifies a arrangement for performing the above method.

The arrangement carries out automated data acquisition, by means of a means for the same, of forms whose design and information is not known in advance, by input into the said arrangement together with storage of patterns of the same. It learns adaptively and registers the design of the form, and includes a computer with the following means for carrying out the adaptive registration:

means for generation of a form map based on the previously unknown form's design for identifying information contained on the form;

means for searching and comparing the form map with stored, registered maps in a means for storing form maps;

means for storing generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

means for indicating agreement according to the limits for agreement when agreement is found; and

means for identification and continued data acquisition of the information content of the form.

In addition, the arrangement can include or constitute that specified according to the above method of the present invention, which is further illustrated in the accompanying non-independent claims for the arrangement.

Brief description of the drawings

Further reference to the enclosed figures and associated text will give a clearer understanding of the present invention.

Fig. I shows schematically how a line pattern is accomplished from a scanned-in invoice.

Fig. 2 shows schematically a flow-path for scanning, identifying, interpreting and validating a form according to the present invention.

Detailed description of preferred embodiments

In the continued description of the present invention, the forms are presented as invoices. However, the invention is not limited to invoice forms but also covers general documents containing text, figures, etc. as forms. Invoices are used here as an example of a form to exemplify the invention.

Fig. 1 illustrates schematically one part of invoice 10 that is scanned in a computer and that is shown on the display. As is evident from invoice 10, it is unclear or blurred after the scanning or input.

Invoice 10 consists partly of a logotype 12 and the vertical 14 and horizontal line 15 elements.

Note that even the logotype contains long black or varying degrees of shaded coloured line elements 16 that have been partly registered in a line map 18 according to the present invention, and that give an idea of what the original logotype 12 looked like, which simplifies identification when the invoice is an object to be identified as being as registered in a form map database. Coloured lines also include grey scales of black.

The form map that in this case constitutes line map 18 has been filtered from other objects 19, such as whole or parts of text objects or coloured objects, plus even the said line elements that include colour, which cannot be reproduced here, but that can be included as many coloured fields on a form 10.

An invoice 10 that is prepared according to the present invention, hereafter designated EH (Eyes & Hands), must be identified at an early stage. For successful identification, EH must have on a previous occasion, learned what the current invoice 10 looks like, which in simple terms means that information about the invoice is available in the form database in EH.

By necessity, the identification must be quick and be able to be made in a database that holds a very large number of invoice identities 18. It is not uncommon for databases to contain more than 10,000 identities 18.

The method and arrangement that EH uses does not require that an invoice is always put through a scanner in exactly the same way, i.e. the information on the invoice can vary somewhat in the x and y axes within a pre-determined measurement or threshold value. Fig. 1 shows a schematic cartographic system of co-ordinates.

In the present invention, (EH) comprises that in one embodiment, EH searches for all vertical 14 and horizontal line 15 elements of a pre-determined length on the invoice. Lines 14, 15 do not need to be free-standing and isolated, but can, for example, be part of a larger logotype text 12, such as ReadSoft AB in Fig. 1. The logotype 12 is represented as the line element 16 in line map 18.

The horizontal lines 15 and the vertical lines 14 constitute the basis for generating a horizontal key (h-key) and a vertical key (v-key) respectively according to the following:

* The invoice is divided into a large number of horizontal segments along the y-axis (not shown). Each segment is equivalent to one position in the h-key. If a certain segment includes one or more line elements 15, a mark or tag is placed in the equivalent key position. If not, an empty space, an inverted mark or anything else that differentiates itself from a mark is used instead.

* A v-key for the vertical line elements 14 is generated in a similar manner along the x-axis.

* The hand v-keys are given designations and together constitute a line key. Following this, a search is performed, which means that the current line key is compared with line keys for known invoices 10 that exist in EH's database. This comparison takes into account that individual lines or line elements 14, 15 can vary somewhat in position, plus that the total pattern of lines can be displaced somewhat according to suitable pre-determined values in the x and y directions, horizontally and vertically respectively.

* The line keys in the database are sorted according to the number of markings (tags), which are used to make the searching effective.

* The search results in a pre-determined number of probable candidates for the identity of the current invoice 10. All candidates are associated with a factor of merit or a probability that they are the current invoice 10.

* The identity of the invoice is finally confirmed by carrying out an interpretation of that known as the RCG-value (RCG- ReCoGnition). The RCG-value is a value at a given position that is unique for a certain invoice/supplier or other form. Examples of such values are bank giro numbers, post-office giro numbers, invoice numbers, total amounts, etc.

The said segments can, for example, form checked patterns that are fine-screened to varying extents according to the relative need for rapid searching.

The line keys can even be implemented on objects formed wholly or partly of text and colours. These are assigned line keys from an object area list that includes x and y-keys for the object. The object area list can, for example, consist of positions for certain selected objects. The

principles for line maps stated above are even appropriate for objects other than line elements to accomplish identification of forms.

If the line keys are not found in the database, this indicates that the invoice is not known, which results in the new line keys being stored in the database that, in this way, is updated in real time.

If necessary, the operator can, via his computer, manually support the whole or part of the adaptive registration and/or identification of a new form or registered form respectively if several alternative candidates are presented as probable according to the factor of merit.

In addition, the present invention includes an arrangement for performing the method according to the above.

The arrangement performs the automatic data acquisition, by means of a means for the same, of forms whose design and information content is not known in advance, by input into the said means together with storage of patterns of the same. It registers in an adaptive manner and learns the design of forms, and includes a computer with the following means for accomplishing the adaptive registration:

means for generating a form map based on the previously unknown form's design for identifying information contained on the form;

means for searching and comparing the form map with stored, recognised maps in a means for storing form maps;

means for storage of generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

means for indicating agreement according to the limits for agreement when agreement is found; and

means for identification and continued data acquisition of the information content of the form.

The, said means are preferably controlled by computer hardware and software, such as, for example:

A scanner for acquisition of data.

An electronic storage medium (hard disk, CD-ROM, etc.) for the means to store information

Signs, icons, signal generators, etc. for indicating purposes.

Filters and comparitors so that the means can search and compare, as well as filters and registers for identification.

On the whole, the means used in the present invention are well known to a skilled person in the technical field, but the way in which they are co-ordinated to achieve the object of the invention is, however, innovative.

In one embodiment of the present invention with reference to Fig. 2, a schematic flow-path is illustrated to show the scanning, identification, interpretation, and validation of a form according to the present invention.

Fig. 2 is divided by dotted lines into partial areas to clarify the different steps in a method according to the invention, whereby the steps constitute the scanning of the form 200, identification of the form 210, interpretation of the form 220, plus validating the form 230.

The form is scanned 200 into EH, and identification 210 follows. Identification consists of generating a line map 212, or alternatively an object area list, whereby an line key is generated. Following this, form 10 is compared 214 with known keys in the form map database, whereby a conformation of identification is obtained via the RCG-value. The next step includes deciding whether the identification was successful 216 according to the conditions "Yes" or "No". If the decision results in "No", a conditional investigation is made to sec if there are more candidates in the form of line keys 218. In the answer here is "Yes", a loop in the form of 214, 216 and 218 is performed until a successful identification is finally made, or until no further line key candidates are presented 218.

In the case of a successful identification, interpretation 220 of the form then begins by interpreting with the help of the current form map 222, after which validation 230 or evaluation 232 of the fields of the form 10 takes place. As an option, the operator can assist with selection if several alternative fields are found 234.

If the identification 210 is unsuccessful, and no further line keys are presented 218, interpretation 220 is performed in that self-learning with a form definition 224 is accomplished. The form definition consists of a template or a set of rules that describes the common elements of a specific collection of forms, for example, Swedish invoices. Following this, the RCG-value is interpreted 226 and a decision is made 228 whether the current RCG-value can be found in the form database. If the answer is "Yes", a re-interpretation begins 229, followed by a continued interpretation 222 that leads to validation 232.

If, on the other hand, the answer is "No", validation commences 230, 236, after which the form is saved in the form map database with the line key 238. Prior to steps 236, 238, the operator can, if several field alternatives are found, assist with the self-learning process.

The embodiments of the present invention described above are possible embodiments, but are not intended to limit the invention to such, as further embodiments will be evident to a skilled person in the technical area via the drafts of the enclosed claims.

Claims

1. Method for the automatic data acquisition by means of a means for the same, of forms whose design and information content is not known in advance, by input into the said means, together with storage of patterns of the same, wherein the method is adaptive and includes self-learning and registration of the design of forms without initial scanning of a blank form sheet being necessary, whereby it includes the following steps to accomplish the adaptive registration:

generation of a form map based on a previously unknown form's design for identifying information contained on the form;

searching and comparing the form map with stored, registered maps in a means for storing form maps;

storage of generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

indication of agreement according to the limits for agreement when agreement is found; and

continued data acquisition for identifying the information content of the form.

2. Method according to claim 1, wherein the form map consists of an object area list with objects contained in the form.

3. Method according to claim 2, wherein the object comprises colours and wholly or partly of text.

4. Method according to claim 1, wherein the form map constitutes a line map comprising line elements from the form.

5. Method according to claim 4, wherein horizontal lines in the line map are used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position horizontal key.

6. Method according to claim 4, wherein vertical lines in the line map are used to generate a vertical key by dividing the form into a pre-determined number of vertical segments

along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

7. Method according to claim 5, wherein at least one line element that is included in a segment is marked in the equivalent key position, and that segments that lack line elements remain unmarked in the equivalent key position.

8. Method according to claim 4, wherein the horizontal key and/or a vertical key constitute a line key in the line map, whereby, during the said searching, the line key generated is compared with line keys stored in the storage means for verifying agreement.

9. Method according to claim 8, wherein the line keys are sorted in the storage means according to the number of markings.

10. Method according to claim 1, wherein the object's horizontal position in the object area list is used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

11. Method according to claim 1, wherein the object's vertical position in the object area list is used to generate a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

12. Method according to claim 10, wherein at least one object that is included in a segment is marked in the equivalent key position, and that segments that lack objects remain unmarked in the equivalent key position.

13. Method according to claim 10, wherein a horizontal key and/or a vertical key constitute an object key in the object area list, whereby, during the said searching, the object key generated is compared with object keys stored in the storage means for verifying agreement.

14. Method according to claim 13, wherein the object keys are sorted in the storage means according to the number of markings.

15. Method according to claim 1, wherein the searching results in a pre-defined number of requested probable candidates for the currently searched form.

16. Method according to claim 15, wherein an operator can support manually the whole or parts of the adaptive registration or identification of the new form or registered forms respectively if several alternative candidates are found as probabilities according to a factor of merit.

17. Method according to claim 1, wherein the identity of the form is confirmed by the data acquisition of a RCG-value.

18. Arrangement for the automatic data acquisition, by means of a means for the same, of forms whose design and information content is not known in advance, by input into the said means together with storage of patterns of the same, wherein it learns adaptively and registers the design of forms without initial scanning of a blank form sheet being necessary, and includes a computer with the following means for carrying out the adaptive registration:

means for generating a form map based on the previously unknown form's design for identifying information contained on the form;

means for searching and comparing the form map with stored, recognised maps in a means for storing form maps;

means for storage of generated form maps in the storage means when they do not coincide with a stored map according to pre-determined limits for agreement;

means for indicating agreement according to the limits for agreement when agreement is found; and

means for identification and continued data acquisition of the information content of the form.

19. Arrangement according to claim 18 wherein the form map consists of an object area list with objects contained in the form.

20. Arrangement according to claim 19, wherein the object comprises colours and/or wholly or partly of text.

21. Arrangement according to claim 18, wherein the form map constitutes a line map comprising line elements from the form.

22. Arrangement according to claim 21, wherein the horizontal lines in the line map are used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

23. Arrangement according to claim 21, wherein the vertical lines in the line map are used to generate a vertical key by dividing the form into a pre-determined number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

24. Arrangement according to claim 22, wherein at least one line element that is included in a segment is marked in the equivalent key position, and that segments that lack line elements remain unmarked in the equivalent key position, and that segments that lack line elements remain unmarked in the equivalent key position.

25. Arrangement according to claim 22, wherein the horizontal key and a vertical key constitute a line key in the line map, whereby, during the said searching, the line key generated is compared with line keys stored in the storage means for verifying agreement.

26. Arrangement according to claim 25, wherein the line keys are sorted in the storage means according to the number of markings.

27. Arrangement according to claim 18, wherein the object's horizontal position in the object area list is used to generate a horizontal key by dividing the form into a pre-determined number of horizontal segments along the y-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the horizontal key.

28. Arrangement according to claim 18, wherein the object's vertical position in the object area list is used to generate a vertical key by dividing the form into a pre-determined

number of vertical segments along the x-axis in a cartographic system of co-ordinates, whereby each segment is equivalent to a position in the vertical key.

29. Arrangement according to claim 27, wherein at least one object that is included in a segment is marked in the equivalent key position, and that segments that lack objects remain unmarked in the equivalent key position.

30. Arrangement according to claim 27, wherein a horizontal key and/or a vertical key constitute an object key in the object area list, whereby, during the said searching, the object key generated is compared with object keys stored in the storage means for verifying agreement.

31. Arrangement according to claim 30, wherein the object keys are sorted in the storage means according to the number of markings.

32. Arrangement according to claim 18, wherein the searching results in a pre-defined number of requested probable candidates for the currently searched form.

33. Arrangement according to claim 32, wherein an operator can support manually the whole or parts of the adaptive registration or identification of the new form or registered forms respectively if several alternative candidates are found as probabilities according to a factor of merit.

34. Arrangement according to claim 18, wherein the identity of the form is confirmed by the data acquisition of a ReCoGnition (RCG)-value which uniquely identifies a form.

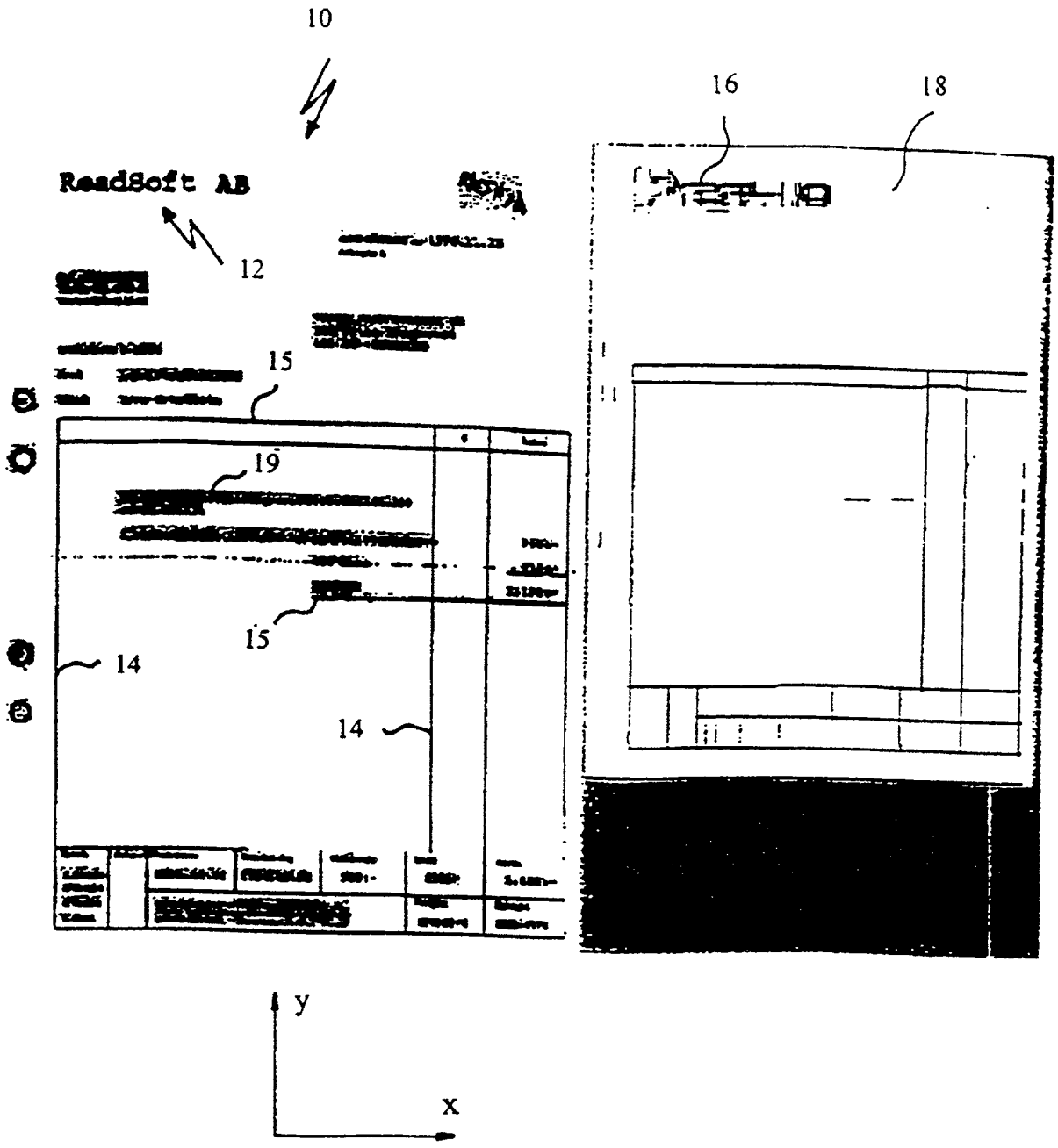

Fig. 1

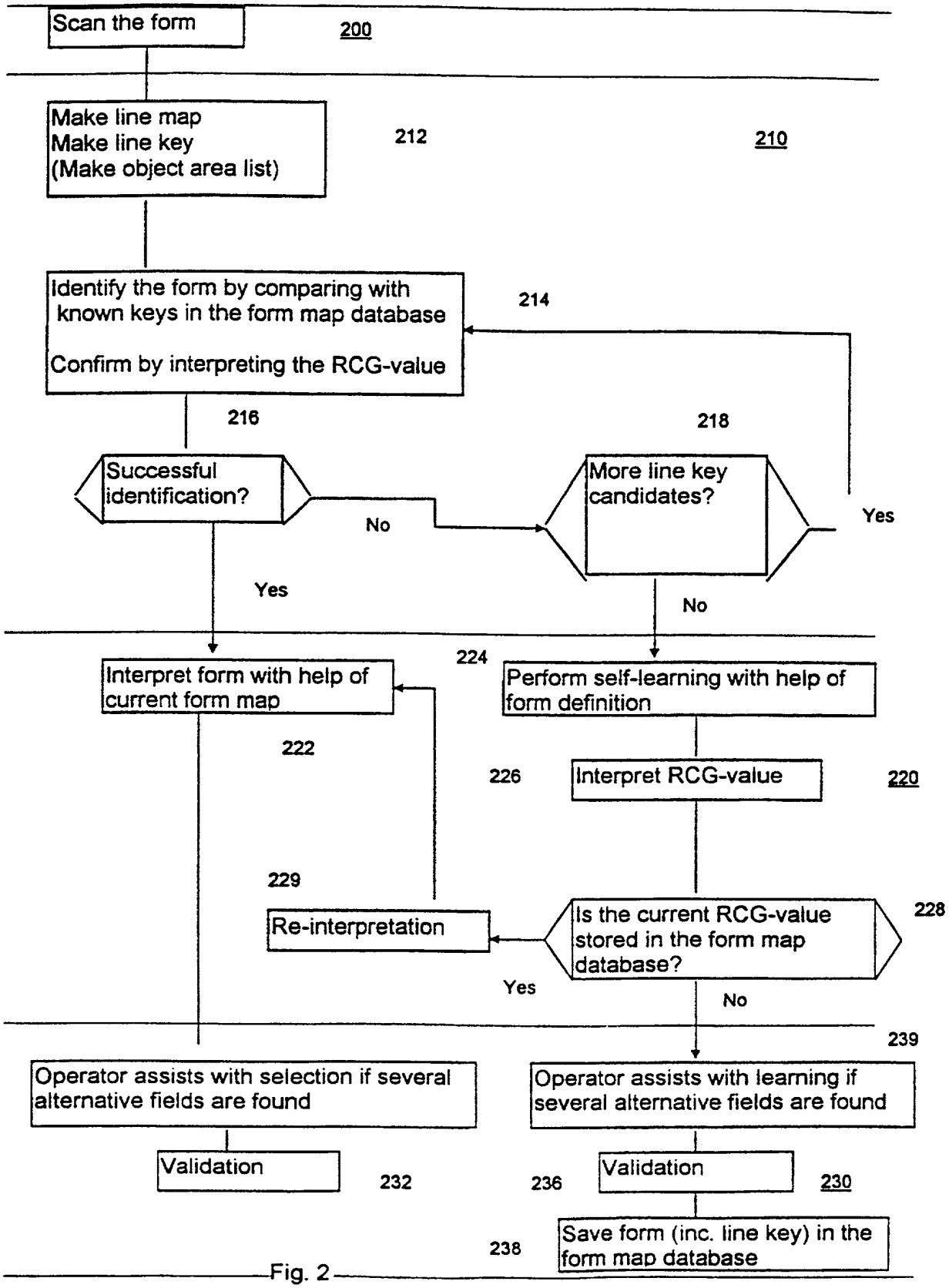
Scan the form
200
Make line map
Make line key
(Make object area list)
212
210
Identify the form by comparing with known keys in the form map database
Confirm by interpreting the RCG-value
214
216
218
Successful identification?
No
More line key candidates?
Yes
Yes
No
224
Interpret form with help of current form map
Perform self-learning with help of form definition
222
226
Interpret RCG-value
220
229
Re-interpretation
Is the current RCG-value stored in the form map database?
228
Yes
No
239
234
Operator assists with selection if several alternative fields are found
Operator assists with learning if several alternative fields are found
Validation
232
236
Validation
230
238
Save form (inc. line key) in the form map database

Fig. 2

09/381899

PTO/PCT Rec'd 16 DEC 1999

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Jan Andersson

Serial No.: 09/381,899 Group No.:

Filed: 04/01/99 Examiner:

For: Method and Arrangement for Automatic Data Acquisition of Forms

Commissioner of Patents and Trademarks
Washington, D.C. 20231

STATEMENT BY ATTORNEY THAT APPLICATION FILED IN PTO IS THE ONE INVENTOR EXECUTED BY SIGNING DECLARATION

NOTE: This form is to be used when the declaration only indicates the name(s) of the inventor(s) and the title of the invention. Notice of September 12, 1983, 1035 O.G. 3.

I, Russell L. McIlwain

Name of Attorney

Jones, Day, Reavis & Pogue

P.O. Address

77 West Wacker Drive, Chicago, IL 60601-1692

Reg. No. 28,641 Tel. No. (312) 782-3939

state I am the attorney for this application and the application identified above is the application which the inventor(s) executed by signing the declaration which is being submitted herewith.

SIGNATURE OF ATTORNEY

CERTIFICATE OF MAILING (37 CFR 1.8a)

I hereby certify that this paper (along with any paper referred to as being attached or enclosed) is being deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to the: Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Merri C. Merrill

(Type or print name of person mailing paper)

Date: 12/10/99

(Signature of person mailing paper)

Please type a plus sign (+) inside this box → [+]

PTO/SB/01 (12-97)
Approved for use through 9/30/00. OMB 0651-0032
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION FOR UTILITY OR DESIGN PATENT APPLICATION (37 CFR 1.63)

☐ Declaration Submitted with Initial Filing **OR** ☒ Declaration Submitted after Initial Filing (surcharge (37 CFR 1.16 (e)) required)

Attorney Docket Number	097037014006
First Named Inventor	ANDERSSON, Jan
COMPLETE IF KNOWN	
Application Number	09 / 381,899
Filing Date	April 1, 1999
Group Art Unit	
Examiner Name	

As a below named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

METHOD AND ARRANGEMENT FOR AUTOMATIC DATA ACQUISITION OF FORMS

(Title of the Invention)

the specification of which
☐ is attached hereto
OR
☒ was filed on (MM/DD/YYYY) 04/01/99 as United States Application Number or PCT International Application Number 09/381,899 and was amended on (MM/DD/YYYY) _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment specifically referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. 119(a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate, or of any PCT international application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application Number(s)	Country	Foreign Filing Date (MM/DD/YYYY)	Priority Not Claimed	Certified Copy Attached? YES	Certified Copy Attached? NO
9701183-7	Sweden	04/01/1997	☐	☐	☒
PCT/SE98/00602	WO	04/01/1998	☐	☐	☒
			☐	☐	☐
			☐	☐	☐

☐ Additional foreign application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto:

I hereby claim the benefit under 35 U.S.C. 119(e) of any United States provisional application(s) listed below.

Application Number(s)	Filing Date (MM/DD/YYYY)

☐ Additional provisional application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

Please type a plus sign (+) inside this box → [+]

PTO/SB/01 (12-97)
Approved for use through 9/30/00. OMB 0651-0032
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

DECLARATION — Utility or Design Patent Application

I hereby claim the benefit under 35 U.S.C. 120 of any United States application(s), or 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. 112, I acknowledge the duty to disclose information which is material to patentability as defined in 37 CFR 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

U.S. Parent Application or PCT Parent Number	Parent Filing Date (MM/DD/YYYY)	Parent Patent Number (if applicable)
PCT/SE98/00602	04/01/1998	

☐ Additional U.S. or PCT international application numbers are listed on a supplemental priority data sheet PTO/SB/02B attached hereto.

As a named inventor, I hereby appoint the following registered practitioner(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: ☐ Customer Number [ ] → Place Customer Number Bar Code Label here
OR
☒ Registered practitioner(s) name/registration number listed below

Name	Registration Number	Name	Registration Number
James B. Raden	~~24,594~~	Marc S. Blackman	~~43,501~~
Joseph H. Golant	~~24,210~~	Alfred E. Hall	~~24,009~~
Ronald A. Sandler	~~23,066~~	Lester Savit	~~30,523~~
Russell L. McIlwain	28,641	Sandra Weiss	30,814

(10)

☒ Additional registered practitioner(s) named on supplemental Registered Practitioner Information sheet PTO/SB/02C attached hereto.

Direct all correspondence to: ☐ Customer Number or Bar Code Label [ ] OR ☒ Correspondence address below

Name	Ronald A. Sandler				
Address	Jones, Day, Reavis & Pogue				
Address	77 West Wacker Drive				
City	Chicago	State	IL	ZIP	60601-1692
Country	U.S.A.	Telephone	(312) 782-3939	Fax	(312) 782-8585

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001 and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Name of Sole or First Inventor: ☐ A petition has been filed for this unsigned inventor

Given Name (first and middle [if any])	Family Name or Surname
Jan	Andersson

Inventor's Signature						Date	99.10.12
Residence: City	Banckagatan 17	State	Viken	Country	Sweden	Citizenship	Swedish
Post Office Address	Banckagatan 17						
Post Office Address							
City	Viken	State		ZIP	SE-254 66	Country	SWEDEN

☐ Additional inventors are being named on the ____ supplemental Additional Inventor(s) sheet(s) PTO/SB/02A attached hereto

Please type a plus sign (+) inside this box → [+]

PTO/SB/02C (3-97)
Approved for use through 9/30/98. OMB 0651-0032
Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

DECLARATION

REGISTERED PRACTITIONER INFORMATION (Supplemental Sheet)

Name	Registration Number	Name	Registration Number
David L. Witcoff	31,443		
John Marlott	37,031		